ASEC 2025 Conference

The 6th International Electronic Conference on Applied Sciences

MDPI

09–11 December 2025 | Online

Deep Learning and Transfer Learning Models of Indian Food Classification

Dr.Jigarkumar Ambalal Patel, Dr.Dileep Laxmansinh Labana, Dr.Gaurang Vinodray Lakhani, Dr.Rashmika Ketan Vaghela

INTRODUCTION & AIM

Food provides an important lens for understanding everyday Indian culture, identity, and social behaviour. In India, eating habits are shaped by factors such as caste, class, religion, family traditions, region, and personal lifestyle. What people choose to eat, how they prepare it, and the occasions on which food is consumed reveal deeper social patterns and emotional connections. With India's expanding economy and growing middle class, food has become a visible part of cultural expression. The country's diverse communities and traditions create a rich and varied food landscape influenced by both local practices and global trends.

Today, rising obesity highlights the need for better calorie awareness. Food-image analysis offers a useful way to support healthier eating and daily meal planning. Recognizing food through images is challenging because dishes vary greatly in ingredients, preparation styles, and presentation, but the availability of food photos from social media and apps has made dataset creation easier. Automated food-recognition systems help users track calories, organize food images, and build healthier habits, while also supporting research and dietary guideline development. Advances in computer vision and machine learning have accelerated progress in this field. Image-assisted and fully image-based dietary assessment methods both show promise, with automated approaches significantly reducing manual effort for nutrition experts.

METHOD

The main aim of this research is to design and evaluate deep learning models capable of accurately classifying Indian food images. Several architectures are explored for their ability to learn visual patterns and distinguish between dishes with diverse textures, shapes, and ingredients. Convolutional Neural Networks (CNNs) form the foundational model, as they are highly effective at extracting spatial and textural features through convolution and pooling layers. These learned features are later passed to fully connected layers for final classification.

DenseNet121 offers an advanced alternative with its densely connected layers, which promote feature reuse and help resolve vanishing gradient issues. Its deep structure allows the model to capture finer details present in Indian dishes. InceptionV3, another model examined in this study, uses parallel convolution filters of different sizes to capture multi-scale features. This ability is particularly useful because Indian food items often have complex patterns at both small and large scales.

MobileNetV2 and MobileNetV3 provide lightweight solutions suited for mobile or edge devices. Their use of depthwise separable convolutions reduces computation while maintaining good accuracy, making them suitable for real-time applications. VGG16, known for its straightforward and uniform design, remains a strong baseline despite its heavier parameter count. Xception enhances Inception's design by fully adopting depthwise separable convolutions, enabling efficient learning of both local and global features—an advantage when classifying visually varied food items.

The research process involves collecting and preprocessing a diverse dataset of Indian food images, followed by adapting each model's final layers to match the number of target categories. Transfer learning with ImageNet-trained weights speeds up training and improves accuracy. Each model is trained and validated with tuned hyperparameters, then evaluated using metrics such as accuracy, precision, recall, and F1-score. Comparative analysis of these results helps identify the most effective architecture for Indian food classification.

RESULTS & DISCUSSION

In this study, we examined the effectiveness of deep learning and transfer learning models for classifying Indian food images. The dataset consisted of 13 food categories collected from Google, offering diverse visual samples but limited quantities per class. Due to this constraint, strategies like data augmentation and pre-trained architectures became crucial for improving model generalization and performance.

A standard CNN model was trained first, achieving strong training accuracy but poor validation accuracy, indicating overfitting. Its overall performance remained moderate, with 50% accuracy and similar precision, recall, and F1-score values. In contrast, transfer learning models—DenseNet121, Inception, MobileNet, VGG16, and Xception—demonstrated significantly better results. DenseNet, Inception, and Xception consistently achieved around 92% across all performance metrics. MobileNet and VGG16 reaching approximately 86–91% accuracy depending on the metric. These results highlight the advantage of using pre-trained architectures for image-based classification tasks, especially when working with limited datasets. Overall, the study shows that transfer learning models, particularly DenseNet, Inception, and Xception, offer robust and reliable performance for Indian food classification.

Model	Accuracy	Precision	Recall	F1-Score
CNN	50%	47%	47%	47%
DenseNet121	92%	92%	91%	91%
Inception	92%	92%	91%	92%
MobileNet	91%	91%	92%	91%
VGG16	86%	86%	85%	86%
Xception	92%	92%	91%	92%

CONCLUSION

This study evaluated Indian food classification using a custom CNN and several transfer learning models, including DenseNet, Inception, MobileNet, VGG16, and Xception. Performance was assessed using accuracy, precision, recall, and F1-score. The CNN achieved moderate results with around 50% accuracy, while transfer learning models performed significantly better. DenseNet, Inception, and Xception achieved 92% across all metrics, and MobileNet and VGG16 reached 86–91%. These results highlight the advantages of pre-trained models, offering superior feature extraction and improved accuracy for diverse Indian food images.

FUTURE WORK

Future work could expand the dataset with diverse Indian food images, use advanced models like EfficientNet or Vision Transformers, explore multi-modal approaches, optimize lightweight models for real-time mobile use, and integrate automated nutritional analysis for personalized diet tracking.